

Silicone Breast Implants: Epidemiological Evidence of Sequelae

Zena A. Stein, MA, MB, BCh

The essential scientific evidence to establish ill effects of silicone breast implants in the body would have to be epidemiological: Does the presence of silicone in the body, perhaps years after implants have been placed in the breast, induce or at least favor the development of distant systemic effects? Angell's review of the regulatory and legal aspects of the silicone breast implant cases makes it immediately apparent that no epidemiological evidence was available when the first important decisions about litigation and compensation were made.¹

The systemic effects first suspected included some form of connective tissue disease and were presumed in turn to reflect autoimmune disease responses. Later, when descriptions of the symptoms reportedly affecting exposed women defied any unitary disease categorization, suppositions about systemic effects turned to less well-defined forms of connective tissue disease that were neither specific nor consistent in expression.

Suspicions of systemic effects originated in Japan in the early 1960s; however, these effects related not to implants but to injections of saline or silicone.² The first US case of compensation awarded to a woman in connection with a silicone implant was in 1990. The plaintiff was awarded about \$1 million; according to Angell, the extant records of this case do not even describe her actual complaints.¹ In 1991, a second litigant who had had an implant (and who received \$7.34 million) was diagnosed as having an autoimmune disease associated with mixed connective tissue disease.¹ (The physician attending the plaintiff apparently gave evidence at the trial that her symptoms preceded the date of the implant.) A third woman, who a year later was awarded \$25 million, said that her symptoms were like "a bad case of flu all the time: with sinusitis, sore throats and colds."¹

Not all readers will fully appreciate the differences between lawyers' and epidemiologists' methods of reaching decisions as to cause. Some of these differences are set out by Annas in this issue³; he does not mention, however, the attention given to diagnostic specificity in the silicone-related trials.

For the epidemiologist, outcomes in the form of clinical diseases might include diverse disorders, but each disorder must, as a minimum requirement, be consistently diagnosed and hence definable on replication. In a recent class suit for a very large group of women, 10 forms of connective tissue disease

were listed as likely effects of silicone. Understandably, these conditions formed the basis of the epidemiological inquiries. The legal trials, by contrast, postulated pathological mechanisms as the hypothetical consequences of silicone exposure, focusing on the formation of autoantibodies and speculating about the processes involved.

One difference discussed by Annas is the weight that epidemiologists and lawyers attached to different levels of probability.³ According to Annas, the cutoff point for a lawyer deciding that a relationship is or is not present requires a simple balance between "is" and "is not": more than half the time or less than half the time. An epidemiologist will want at least a 95% level of certainty, preferably with a confidence limit that does not include 1.0.

But the epidemiologist, by training and experience, is influenced not only by probability but also (heavily) by the research design and rigor yielding the probabilities. Thus, a clinician may report a single case or a collected series of manifestations and, in a published report, legitimately speculate on a possible cause.⁴ This speculation, traditionally respected by epidemiologists as the "alert clinician" observation, will often trigger epidemiological investigations. Often, the next step will be a quick and low-cost case-control study, of which there are many among the silicone implant studies (e.g., Strom et al.⁵ and Hochberg et al.⁶). In instances in which a disease is rare (as are many connective tissue diseases) and exposure relatively common, this step is a natural starting point for investigating the suspected cause-effect relationship.

However, because in the case of silicone the clinicians reported not just one but several different diseases, a case-control design might not be completely convincing; also, and worse, some of the supposed aftereffects seem not to be clear-cut syndromes or conditions that were previously regarded as diagnostic entities. Shifting targets are not suitable for case-control studies.

Cohort designs are preferable on this account. Such designs theoretically start with the exposed population (here, those with a silicone implant) and follow these individuals, comparing them on a full range of their diseases and health disorders with carefully chosen, perhaps matched comparison, populations. Cohort designs were selected for detailed discussion here.

The Epidemiological Evidence

Although the epidemiological studies got off to a late start, we can now move well beyond the early case reports and examine the considerable evidence on the possibility of an association between well-described forms of connective tissue disease and silicone implants. First, there are well-conducted published case-control studies of the association of subacute lupus erythematosus⁵ and scleroderma,⁶ both very rare conditions, with silicone implants. Controls should be as similar to case patients as possible, in all respects except those that might increase or decrease history of implant surgery. These studies and other similar investigations have yielded odds ratios of around 1.0, indicating that women with the specific (often rare) disorder were no more likely than women who did not have the disorder to report a history of implant surgery.

Sound and well-executed cohort studies that can explore multiple end points, described in different ways, better clarify the particular issues at stake. The 5 largest are described subsequently, and the results are summarized in Table 1.⁷⁻¹¹

Study 1

The 5 cohort studies are all different from each other in certain ways, and loopholes, if aggressively sought, can be found or argued in each. Together, because they replicate and supplement each other in important ways, they add up to weighty evidence. The first to appear, in the *New England Journal of Medicine* in 1994,⁷ is the smallest but perhaps the most difficult to fault, and it provides a good starting point.

This cohort emanated from Olmsted County, the population that has been served by the Mayo Clinic and associated physicians since 1907. The study made use of the systematic patient data maintained by the clinic. All physician visits, all investigations, and all procedures relating to each client were placed on a unit record. Thus, it was not difficult to identify the 749 women who had had breast implant surgery, including cases involving silicone, over the 30 years preceding the study.

Editor's Note. See related articles by Macklin (p 487), Annas (p 490), and Fox (p 493) in this issue.

TABLE 1—Relative Risks for Connective Tissue Disorders and Other Categories of Disorders Following Breast Implant Surgery: 5 Cohorts

	Mayo Clinic Study ⁶	Nurses' Health Study ⁷	Women's Health Study ⁸	Danish Hospital Discharges ⁹	Swedish Hospital Discharges ¹⁰
No. of breast implant subjects	749 ^a	876 ^a	10 830 (type not known)	2570 ^a	7442 ^a
Any connective disease, RR (95% CI)	1.06 (0.34, 2.97)	0.3 (0.0, 1.9)	1.24 (1.06, 1.41) ^{a**}	1.1 (0.2, 3.4)	1.1 (0.8, 1.6)
Any arthritis, RR (95% CI)	1.35 (0.81, 2.23)
Signs/symptoms, documented, RR (95% CI)	...	0.6 (0.2, 1.6)	...	2.5 (1.7, 3.5)	...
Signs/symptoms, self-reported, RR (95% CI)	...	1.2 (0.7, 2.2)
Other, including mixed, self-reported, RR (95% CI)	1.30 (1.06, 1.62)*

^aSilicon gel filled.

*P = .017; **P = .0015.

The complete records for each of these patients over a mean of 10 years following surgery were extracted. For comparison, additional records of 1948 patients, 2 women matched by age to each woman "exposed" to implant surgery, were selected. Each had had at least 1 medical evaluation documented within 2 years of the date of implant of the exposed woman. Trained nurses extracted all information relating to a full range of connective tissue disorders, as well as others in a spectrum of disorders with autoimmune origins (e.g., Hashimoto's disease), cancers other than breast cancer, sarcoidosis, and "any arthritis." The extractors were carefully monitored, and the team examined each record in detail. Less serious as well as major complaints were also included in the record study (e.g., "morning stiffness").

The incidence rate of 8 per 10 000 person-years for all connective tissue disorders diagnosed by the end of 1991 (i.e., before publicity about silicone erupted) was virtually identical for exposed and unexposed women. The equality in rates continued throughout the list of suspected disorders, signs, and symptoms. The sole exception was morning stiffness, and even this condition was confined to the minority of women whose implant followed mastectomy for cancer. Not a single case of morning stiffness was recorded for "augmentation" implants.

The equality of calculated risks for these rare diseases allowed, as always, for a range around the estimated values, which in this instance was less than 2 for every individual diagnosis. Thus, it is very unlikely that any excess would exceed a doubling in rate. The consistency of the results for so large a list of conditions is remarkable, and it is hard to fault this result, except to note that the sample size was insufficient to exclude a "small" effect and therefore the possibility that a small subgroup of women might be affected.

Some have observed that research support for the study came not only from the National Institutes of Health but also, in part, from the Plastic Surgery Educational Foundation. This support is openly acknowledged, however, and I see nothing to substantiate a charge of conscious or unconscious bias. The single apparent source of potential bias involves the "unblinded" record abstractors, an element difficult and probably unnecessary to avoid in this type of study.

Study 2

Sound epidemiological inference requires replication to attain the level of certainty required for important therapeutic or preventive actions, and, in a surprisingly short time, the Mayo Clinic publication was followed in the *New England Journal of Medicine* by a second report, this time from the Harvard Nurses' Health Study.⁸ In the Harvard study, each participant is asked to complete a biennial health questionnaire. Since 1980, the questionnaire has included items on the connective tissue diseases diagnosed since 1976; subacute lupus erythematosus was mentioned specifically on the questionnaire in 1982, 1984, and 1992; rheumatoid arthritis has been included since 1982; and scleroderma, polymyositis, dermatomyositis, and Sjögren's disease have been included since 1992. The item "other major illness diagnosed" was repeated on every biennial questionnaire.

In 1992, a supplementary questionnaire was sent to all participants who had responded positively to any of the items just described. For all participants who had previously noted having a connective tissue disorder, the diagnosis was then confirmed or disconfirmed on the basis of medical records from 1976 to 1990.

This procedure yielded 3 overlapping groups of respondents: 516 subjects with a confirmed diagnosis on the record, 904 for whom the record indicated signs or symptoms that might indicate connective disease but no particular syndrome or a minor complaint, and 1305 who personally reported signs or symptoms of connective tissue disease.

Of the respondents, 1183 reported that they had had breast implant surgery. This information was also validated: medical records for a random sample of 100 respondents corresponded almost completely with respondent reports.

Analyses were based on person-years of exposure. Exposure levels were analyzed separately for each of the categories of certainty just described (and shown in Table 1). A large number of women with probable or minor disorders were also studied. Relative risks for definite and suspected diseases among women with breast implants (whether silicone filled or not), as compared with women without implants, were all around 1.0; in no case did the upper confidence limit exceed 2.2.

As with the Mayo Clinic study, one can exclude, from a statistical perspective, a large effect (e.g., a doubling of risk). Were the 2 studies to be examined side by side, it is difficult to believe they could be hiding a large effect, at least one involving the diseases mentioned. Furthermore, there seemed to be no large effect for less severe or less well-defined syndromes resembling connective tissue disorders.

The Mayo Clinic investigation was reconstructed from prospectively recorded data, and none of the patients had to be excluded from the study; the Harvard Nurses' Health Study attempted to block sources of bias or omission that would confound the results. Each study has allowed for a long postsurgery period, at least 10 years.

In both instances, the study under discussion was only one of many undertaken using the particular data sets, adding to the confidence that can be placed in the results.

Study 3

Results of the third cohort study, again involving data collected for other purposes, were published in the *Journal of the American Medical Association*.⁹ This was neither a clearly prospective study, as exemplified by the Mayo Clinic example, nor a retrospective cohort study, like the Nurses' Health Study (in which the outcomes were recorded in part when they occurred but also retrospectively, when the special questionnaire asked, for example, about breast implant history).

The population used for the third study was self selected from 1.75 million women in the health professions who received mailed questionnaires. Some 400 000 of these women volunteered to participate in the Women's Health Study and filled out a detailed questionnaire asking about a range of health experiences. A total of 805 respondents reported connective tissue disease between 1962 and 1991. At the end of the questionnaire were questions about breast implants; 10830 women reported having had implants. Although respondents were asked to indicate the date of onset of symptoms and the date of the surgery, there were many missing items for which imputation was needed in the analysis.

The mailing took place between 1992 and 1996, at the height of concern related to possible associations between implants and connective tissue disease. There was no confirmation of the fact, date, or type of breast surgery, and self-reported information on connective tissue disease was not validated. Although respondents were all connected in some way with the health professions, this may be insufficient grounds to infer that the diagnoses, if made, were remembered or understood by the respondents. Neither errors of omission nor errors of commission can be excluded.

The rates for the 6 disorders among women exposed to breast implants closely resembled those found in the other 2 series; all rates were close to 1.0. The relative risk for "any one of the connective diseases listed" was 1.24, which, in this very large series, was statistically significant; however, the upper 95% confidence limit was 1.4, suggesting that while there may have been a slight excess among those reporting implants, a doubling of risk would be unlikely.

Even the single slight but statistically significant finding for these self-reported conditions must be considered against the

climate of the period during which the study was conducted, when litigation was at its height and related news reports reached a wide audience. Aspects of this study are to be repeated in an attempt to validate the presence of the reported diseases and disorders.

Studies 4 and 5

Two further studies, one from Denmark¹⁰ (published in 1997) and the other from Sweden¹¹ (published in 1998), carry additional weight based on populations somewhat less affected by the US media. Each of these studies involved national hospital discharge registries of both breast implant surgeries and connective tissue disorders. Individual records were linked, and, as shown in Table 1, considerable numbers of respondents were exposed to surgery. Again, the selective risks for connective tissue diseases—1.1 in each study—were almost identical for the exposed and unexposed groups. Large numbers of respondents and detailed medical notes lent strength to these studies. On the other hand, because only patients with severe cases were admitted for treatment, patients with less severe cases would have been omitted from the studies.

Summary

Skeptics may certainly find fault with the third study (the only one to report a significant finding) or with all or any of the statistics described. But few could argue, after examining these studies, that the relative risk for a known and well-defined connective tissue disease is likely greater than 2.

Another possibility has been raised, one that calls to mind other difficult-to-study syndromes linked to exposures. In a study addressing Gulf War syndrome,¹² signs and symptoms were often mentioned that proved difficult to describe systematically and therefore difficult to study. What if a suspected silicone exposure syndrome were so mild and transient that it did not warrant a physician visit (as in the Mayo Clinic study), receive a diagnosis (as in the Nurses' Health Study), or require admission to a hospital (as in the Danish and Swedish studies)? And if such a disorder were, in fact, mild and transient, should it merit the concern that has been shown and the compensation that has been awarded in the silicone implant litigation cases to date? We suggest that neither a well-described disorder with a relative risk of less than 2 nor a transient and mild disorder seems compatible with the number of litigants over silicone implants and the apparent seriousness of their complaints.

Some 400 000 women¹ joined in one class action suit for damages, and 170 000 joined in another.¹³ Even if there had been 2 million implants undertaken in the United States over the 3 decades in which implant surgery has been practiced (and some estimates put the number closer to 1 million¹⁴), there is no conceivable way in which a relative risk of 2 or 3 (or even 4) for each of the rare syndromes reported could explain so many exposed women being affected. At most, 2200 out of 2 million unexposed women would be expected to have had any one of the listed forms of connective tissue disorders, most of which are very rare. Doubling the risk among the exposed population yields 4400, and increasing the risk 20-fold produces 44 000. At this rate, there is no way in which 400 000 litigants could all be affected.

Extensions of the already-completed studies are ongoing,^{15,16} at least 1 of which is government funded; apparently it is thought in the United States (though not in the United Kingdom¹⁷ or elsewhere) that there is still room for reasonable doubt as to the supposed causal relationships. But if epidemiology is invoked in the interest of public health to prevent the many uses of silicone, the weight of the evidence abstracted here supports the inference that silicone breast implants have not been proved guilty of causing connective tissue disorders. □

References

1. Angell M. *Science on Trial: The Clash of Medical Evidence and the Law in the Breast Implant Case*. New York, NY: WW Norton & Co; 1996.
2. Miyoshi K, Miraoko T, Kobayashi Y, et al. Hypergammaglobulinemia by prolonged adjuvanticity in man: disorders developed after augmentation mammoplasty. *Jpn Med J*. 1964;2122:9-14.
3. Annas GJ. Burden of proof: judging science and protecting public health in (and out of) the courtroom. *Am J Public Health*. 1999;89:490-493.
4. Vasey FB, Havice DL, Bocanegra TS, et al. Clinical findings in symptomatic women with silicone breast implants. *Semin Arthritis Rheum*. 1994;24:22-28.
5. Strom BL, Reidenberg MM, Freundlich B, Schinnar R. Breast implants and risk of system lupus erythematosus. *J Clin Epidemiol*. 1994; 47:1211.
6. Hochberg MC, Perlmutter DL, Medsger TA, et al. Lack of association between augmentation mammoplasty and systemic sclerosis (scleroderma). *Arthritis Rheum*. 1996;39:1125-1131.
7. Gabriel SE, O'Fallon WM, Kurland LT, Beard CM, Woods JE, Melton LJ III. Risk of connective-tissue diseases and other disorders after breast implantation. *N Engl J Med*. 1994;330:1697-1702.
8. Sanchez-Guerrero J, Colditz GA, Karlson EW, Hunter DJ, Speizer FE, Liang MH. Silicone breast

- implants and the risk of connective tissue diseases and symptoms. *N Engl J Med*. 1995;332:166–170.
9. Hennekens CH, Lee I-M, Cook NR, et al. Self-reported breast implants and connective-tissue diseases in female health professionals. *JAMA*. 1986;275:516–521.
 10. Fris S, Mellemkjær L, McLaughlin JK, et al. Connective tissue disease and other rheumatic conditions following breast implants in Denmark. *Ann Plast Surg*. 1997;39:1–8.
 11. Nvrén O, Yin L, Josefsson S, et al. Risk of connective tissue disease and related disorders among women with breast implants: a nation-wide retrospective cohort study in Sweden. *BMJ*. 1998; 316:417–422.
 12. Wegman DH, Woods NF, Bailar JC. Invited commentary: How would we know a Gulf War syndrome if we saw one? *Am J Epidemiol*. 1997;146:704–711.
 13. Stolberg SG. Neutral experts begin studying dispute over breast implants. *New York Times*. July 23, 1998:A21.
 14. Terry MB, Skovron ML, Garber S, Sorenson E, Toniolo P. The estimated frequency of cosmetic breast augmentation among US women, 1963 through 1988. *Am J PublicHealth*. 1995;85:1122–1124.
 15. Jung B. Scientific and medical care issues still unresolved in breast implant controversy: reports and studies eagerly awaited. *Epidemiol Monitor*. 1998;19:5.
 16. Brown SL, Langone JJ, Brinton LA. Silicone breast implants and autoimmune disease. *J Autoimmun Dis Womens Health*. 1998;53:21–24.
 17. *Silicone Gel Breast Implants: Report of the Independent Review Group*. Cambridge, England: Jill Rogers Associates; 1998.

Ethics, Epidemiology, and Law: The Case of Silicone Breast Implants

Ruth Macklin, PhD

Epidemiology is a powerful tool for reaching scientific conclusions. Like virtually every method science uses to draw conclusions, it is probabilistic. As epidemiologists have continued to refine their methodology, experts have come to agree on what level of certainty must be reached to draw a conclusion that a causal connection exists between 2 (or more) events or conditions. As Zena Stein reminds us in her article,¹ the epidemiological evidence relating to silicone breast implants has been insufficient to demonstrate that the implants caused the systemic effects. The conclusions of previous studies, as in virtually all of science, have been probabilistic; that is, they have demonstrated not that the implants could not have caused the effects, only that it was highly unlikely that they did. Stein contends that “the essential scientific evidence to establish ill effects of silicone in the body would have to be epidemiological.”^{1(p484)} This is the case, however, only in the absence of a strong scientific hypothesis that could provide a plausible causal explanation of the autoimmune symptoms in a subset of all of the women who manifested such symptoms. Were such a hypothesis to be forthcoming, it could be consistent with the epidemiological evidence showing that the implants did not cause the symptoms in all of the women.

George Annas contrasts the process by which conclusions are reached in epidemiological studies with what goes on in the courtroom.² Annas cites the 1993 Supreme Court decision in *Daubert v Merrell Dow Pharmaceuticals*: “Scientific conclusions are subject to perpetual revision. Law, on the other hand, must resolve disputes finally and quickly.”^{2(p491)} The law can and often does resolve disputes finally and quickly, but—at least sometimes—not without a sacrifice of truth or accuracy. Just as scientific hypothe-

ses are often revised or even overturned, so, too, are judicial decisions sometimes reversed.

Can ethics help to illuminate the debate between those who would defend the settlement that promised to compensate almost 200 000 women for breast implant-related injuries and those who criticize the legal settlement on scientific grounds? There is, in this case, little doubt that most of the women manifested clinical signs of the symptoms they claimed were caused by the implants. If, against the probabilistic odds, silicone breast implants actually did harm the women who received them, it would be unjust not to compensate these women. On the other hand, if the epidemiological evidence showing lack of causation reflected the reality of the situation, then the company's expense in a settlement of \$3 billion would be unjust.

One ethical question can thus be framed in terms of justice. Which outcome would be more unjust: failure to compensate the women if the implants did, in fact, cause systemic disease or throwing a pharmaceutical company into bankruptcy for unwarranted product liability? Reasonable people are likely to disagree in response to this question. The women would justly deserve compensation for ill health, disability, or chronic discomfort. The company would not deserve the loss of billions of dollars and the need to radically restructure. Yet, it is hard to escape the conclusion that human beings, afflicted with symptoms ranging from mild discomfort to serious connective tissue or autoimmune disorders, suffer more than a corporation, however great its financial loss.

Our society recognizes other situations involving compensation in which a perfect fit between cause and effect does not exist, yet these arrangements are not considered

unjust. Workman's compensation for injury sustained on the job is one example; no-fault automobile insurance is another. The overall consequences of having these no-fault arrangements are held to be better than the alternative that would require attribution of cause. Although product liability in general does not accept this no-fault scheme, it may be ethically justified in certain cases given the probabilistic nature of epidemiological evidence and the fact that the sample size might not be large enough to detect extremely rare events. For example, the possibility exists that autoimmune diseases affect a subset of genetically susceptible people but that the genetic condition is too rare to be detected by an epidemiological study of this size.

Zena Stein invokes a comparison with Gulf War syndrome, noting that, as with the clinical syndrome of women who had breast implants, the Gulf War symptoms were difficult to describe systematically and thus hard to study. The case of silicone breast implants is made even more difficult by the long latency period between exposure and onset of symptoms. There are similarities and differences between the ongoing controversy over Gulf War syndrome and the situation regarding the breast implants. In the Gulf War case, unlike that of the breast implants, epidemiological evidence has indicated some connection between symptoms (not explained by other medical and psychological conditions) reported by about 20% of the veterans

Requests for reprints should be sent to Ruth Macklin, PhD, Albert Einstein College of Medicine, Department of Epidemiology and Social Medicine, 1300 Morris Park Ave, Bronx, NY 10461.

Editor's Note. See related articles by Stein (p 484), Annas (p 490), and Fox (p 493) in this issue.